# SGA based symbol detection and EM channel estimation for MIMO systems

Yugang Jia*,Christophe Andrieu†, Robert J. Piechocki*, and Magnus Sandell‡
*Centre for Communications Research, University of Bristol, UK.
†Department of Mathematics, University of Bristol, UK.
‡Toshiba Research Europe Ltd., Bristol, UK.
Email: {yugang.jia,c.andrieu,r.j.piechocki}@bristol.ac.uk, Phone: +44 (0) 117 9545 125, Fax: +44 (0) 117 9545 206.

***Abstract*— This paper investigates iterative channel estimation and symbol detection for spatial multiplexing multiple input multiple output (MIMO) systems with frequency flat block fading channels using the expectation-maximization (EM) algorithm. The maximum likelihood (ML) estimation of the MIMO channels via the EM algorithm requires the computation of the posterior mean and covariance of transmit symbol vectors which involve an exhaustive search of all possible symbol combinations and are computationally prohibitive for large systems. However, most of the symbol combinations contribute very little to the estimation. Therefore, we suggest that sequential Gaussian approximation (SGA) algorithm can be used to identify the $M$ most significant symbol combinations and we can approximate the mean and covariance based on those symbol combinations. Simulation results are provided to illustrate the proposed algorithm.**

## I. Introduction

Communication systems with multiple transmit/receiver antennas have been shown to be a very promising technology for the next generation communication systems[1]. To achieve a high performance in such a system, the joint channel estimation and symbol detection plays an important role in the receiver design. In general, the Maximum Likelihood (ML) solution requires prohibitive computation for large systems.

The expectation-maximization (EM) algorithm [2] [3] provides a general framework to approximate the ML solution with reduced complexity in an iterative manner. A variety of EM/SAGE based algorithms have been proposed in the literature [4] [5] [6] [7] [8] [9] to estimate parameters of interest.

The approximate ML estimation of channels via EM algorithm requires the computation of the posterior mean and covariance of the transmitted symbols which involves an exhaustive search of all possible symbol combinations to estimate the joint posterior symbol probabilities and is computationally prohibitive for large systems.

Many suboptimal symbol detection algorithms, such as sphere decoder (SD) [10], successive inference cancelation [7] and probability data association [11], can approximate the marginal symbol probabilities and the mean of transmitted symbols with a reduced complexity. SD type algorithms tend to perform very well but suffer from the fact that their complexity is varying and depends on the channel realization [10].

The joint posterior symbol probabilities are always approximated by the product of the marginal symbol probabilities computed via suboptimal detection algorithms and as a result, the covariance matrix is approximated as diagonal [7] [8].

In this paper, we consider ML channel estimation via EM algorithm for spatial multiplexing MIMO systems with frequency flat block fading channels. Other than approximating the covariance matrix as diagonal, we compute all the elements of the covariance matrix via a suboptimal procedure which does not require searching of all the possible symbol combinations. This is based on the fact that most of the symbol combinations contribute very little to the final result and thus can be approximated with only $M$ significant symbol combinations. The $M$ most significant symbol combinations are identified via sequential Gaussian approximation (SGA) algorithm [12] which has been proposed as a near optimal symbol detector for MIMO systems.

The paper is organized as follows. Section II describes the system model. Section III explains how the EM algorithm works for our system. Section IV illustrates how to identify the $M$ most significant symbol combinations via the SGA symbol detection algorithms and how to compute the mean and covariance based on those symbol combinations. Section V summarizes the proposed algorithm. Section VI demonstrates the performance of the proposed algorithm via Monte Carlo simulation. The paper is concluded in section VII.

## II. System Description

Consider a narrow band spatial multiplexing MIMO systems with $N_T$ transmit antennas and $N_R$ receive antennas. At each time instant $k$, the system model is:

$$\mathbf{y}(k) = \mathbf{H}\mathbf{x}(k) + \mathbf{n}(k) \tag{1}$$

where $\mathbf{H}$ is the $N_R \times N_T$ Rayleigh block fading channel matrix with $h_{(i,j)}$ as its $(i,j)$th entry, which is the channel gain from transmit antenna $j$ to receive antenna $i$; $i = 1,\dots,N_R$ and $j = 1,\dots,N_T$; $\mathbf{x}(k) \stackrel{\text{def}}{=} [x_1(k),\dots,x_{N_T}(k)]^T$ ( $[*]^T$ means transpose, $[*]^*$ means conjugate and $[*]^H$ means conjugate transpose ); a symbol $x_j(k)$ transmitted from the $j$-th antenna is taken from a modulation constellation $A = \{a_1, a_2, \dots, a_N\}$; $\mathbf{n}(k)$ is a $N_R \times 1$ zero-mean complex circularly symmetric Gaussian noise with variance matrix $\sigma_n^2\mathbf{I}$ ($\mathbf{I}$ is the identity matrix).

0-7803-9392-9/06/$20.00 (c) 2006 IEEE

Let $\mathbf{Y}_{1:K} \stackrel{\text{def}}{=} \{\mathbf{y}(1), \ldots, \mathbf{y}(K)\}$ denote all the observations and $\mathbf{X}_{1:K} \stackrel{\text{def}}{=} \{\mathbf{x}(1), \ldots, \mathbf{x}(K)\}$ denote all the transmitted symbols in the current transmission burst, the joint density function for $\{\mathbf{Y}_{1:K}, \mathbf{X}_{1:K}, \mathbf{H}\}$ factors as:

$$p(\mathbf{Y}_{1:K}, \mathbf{X}_{1:K}, \mathbf{H}) = p(\mathbf{H}) \prod_{k=1}^{K} p\Big(\mathbf{y}(k)|\mathbf{x}(k), \mathbf{H}\Big) p(\mathbf{x}(k)) \tag{2}$$

where the probability density function $p(\mathbf{H})$ is the prior knowledge of the channel statistics and $p(\mathbf{x}(k))$ is the prior information about symbols.

Then the log-function will be:

$$\begin{aligned}\log p(\mathbf{Y}_{1:K}, \mathbf{X}_{1:K}, \mathbf{H}) \propto \quad & \sum_k |\mathbf{y}(k) - \mathbf{H}\mathbf{x}(k)|^2/\sigma_n^2 \\ & + \log p(\mathbf{H}) + \sum_k \log p(\mathbf{x}(k)) \end{aligned} \tag{3}$$

## III. ML channel estimation via EM algorithm

The EM algorithm provides an iterative scheme to approach the ML estimate $\hat{\mathbf{H}} = \text{argmax}_{\mathbf{H}} p(\mathbf{Y}_{1:K}|\mathbf{H})$ when the direct computation is prohibitive. The EM algorithm relies on iterations between two steps: a conditional expectation (E) step and a maximization (M) step [3].

The E step finds the expected value of the complete-data log-likelihood $\log\big(p(\mathbf{Y}_{1:K}, \mathbf{X}_{1:K}|\mathbf{H})\big)$ with respect to the unobserved data $\mathbf{X}_{1:K}$ conditioned on the current channel estimate $\mathbf{H}^{(l)}$:

$$\begin{aligned}&\mathcal{Q}\left(\mathbf{H}, \mathbf{H}^{(l)}\right) = E_{\mathbf{X}_{1:K}|\mathbf{Y}_{1:K},\mathbf{H}^{(l)}} \Big(\log\big(p(\mathbf{Y}_{1:K}, \mathbf{X}_{1:K}|\mathbf{H})\big)\Big) \\ &\propto -\frac{1}{\sigma_n^2} \sum_{k=1}^{K} \Big( |\mathbf{y}(k) - \mathbf{H}\bar{\mathbf{x}}^{(l)}(k)|^2 + \text{Tr}\left(\mathbf{H}\mathcal{S}^{(l)}(k)\mathbf{H}^H\right)\Big),\end{aligned} \tag{4}$$

where:

$$\begin{aligned}\bar{\mathbf{x}}^{(l)}(k) &\stackrel{\text{def}}{=} E\big(\mathbf{x}(k)|\mathbf{y}(k), \mathbf{H}^{(l)}\big) \stackrel{\text{def}}{=} [\bar{x}_1^{(l)}(k), \ldots, \bar{x}_{N_T}^{(l)}(k)]^T \\ \bar{x}_j^{(l)}(k) &= \sum_{x_j(k) \in A} x_j(k) p\left(x_j(k)|\mathbf{H}^{(l)}, \mathbf{y}(k)\right)\end{aligned} \tag{5}$$

where $\text{Tr}(*)$ means the trace of a matrix and the matrix $\mathcal{S}^{(l)}(k)$ is the covariance of the posterior symbol probabilities with its $(i,j)$th element defined as follows:

$$\begin{aligned}&\mathcal{S}_{(i,j)}^{(l)}(k) = \\ &E\Big( \left(x_i(k) - \bar{x}_i^{(l)}(k)\right)^H \left(x_j(k) - \bar{x}_j^{(l)}(k)\right) |\mathbf{y}(k), \mathbf{H}^{(l)}\Big)\end{aligned} \tag{6}$$

The channels are updated in the M step by maximizing Eq.(4) with respect to $\mathbf{H}$:

$$\mathbf{H}^{(l+1)} = \mathbf{Y}_{1:K} \left(\bar{\mathbf{X}}_{1:K}^{(l)}\right)^H \left(\bar{\mathbf{X}}_{1:K}^{(l)} \left(\bar{\mathbf{X}}_{1:K}^{(l)}\right)^H + \sum_{k=1}^{K} \mathcal{S}^{(l)}(k)\right)^{-1} \tag{7}$$

with $\bar{\mathbf{X}}_{1:K}^{(l)} \stackrel{\text{def}}{=} [\bar{\mathbf{x}}^{(l)}(1) \ldots \bar{\mathbf{x}}^{(l)}(K)]$.

The computation of $\bar{\mathbf{x}}^{(l)}(k)$ and $\mathcal{S}^{(l)}(k)$ requires the marginal symbol probability $p\left(x_j(k)|\mathbf{H}^{(l)}, \mathbf{y}(k)\right)$ as well as the joint symbol probability $p\left(x_j(k), x_i(k)|\mathbf{H}^{(l)}, \mathbf{y}(k)\right)$ for $i, j = 1, \ldots, N_T$. The optimal computation requires the searching of all possible symbol combinations and is computationally prohibitive for large systems.

However, most of the symbol combinations contribute very little to the estimation. This motivates the following section, where we approximate the mean and covariance based on only $M$ most significant symbol combinations identified via SGA algorithm.

## IV. Mean and covariance approximation via SGA based symbol detection algorithm

In this section, we will review how to identify the $M$ most significant symbol combinations using the SGA algorithm and then illustrate how to compute $\mathcal{S}^{(l)}(k)$ and $\bar{\mathbf{x}}^{(l)}(k)$ based on those symbol combinations.

### A. *M most significant symbol combinations identification*

In the SGA algorithm [12], a suboptimal procedure via Gaussian approximation is proposed to identify the $M$ most significant symbol combinations from all the possible symbol combinations. The main idea of the SGA algorithm is as follows. The selection of $M$ most significant symbol combinations for $N_T$ antennas is decomposed to $N_T$ steps In the $j$-th step ($j = 1, \ldots, N_T$), $M$ most significant symbol combinations are selected for antennas $1, \ldots, j$ based on the $M$ available most significant symbol combinations selected for antennas $1, \ldots, j-1$ and a Gaussian approximation for the additive noise and interference from antennas $j+1, \ldots, N_T$.

Assume that for $k$-th time instant with channel estimation $\mathbf{H}^{(l)}$ and $j \geq 1$, at the $(j-1)$-th step of the algorithm we have identified $M$ significant combinations $\Theta_{j-1}^{(l)}(k) \equiv \{x_1^{(m,l)}(k), \ldots, x_{j-1}^{(m,l)}(k), m = 1, 2, \ldots, M\}$ for antennas $1, 2, \ldots, j-1$. We would like to calculate $p(x_1^{(m,l)}(k), \ldots, x_{j-1}^{(m,l)}(k), x_j^{(l)}(k)|\mathbf{y}(k), \mathbf{H}^{(l)})$ for all $m = 1, \ldots, M$ and $x_j \in A$ in order to select $\Theta_j^{(l)}(k)$ which contains $M$ symbol combinations of the largest probabilities, among the $MN$ possibilities. However this quantity requires prohibitive computations, and instead we choose a Gaussian approximation.

Provided that $(\mathbf{H}^{(l)})^H \mathbf{H}^{(l)}$ is invertible one can rewrite Eq. (1) as follows,

$$\begin{aligned}\tilde{\mathbf{y}}(k) &= \mathbf{x}(k) + \tilde{\mathbf{n}}(k) \\ &= \sum_{q=1}^{j} x_q(k)\mathbf{e}_q + \sum_{q=j+1}^{N_T} x_q(k)\mathbf{e}_q + \tilde{\mathbf{n}}(k) \\ &\stackrel{\text{def}}{=} \sum_{q=1}^{j} x_q(k)\mathbf{e}_q + \hat{\mathbf{n}}_j(k),\end{aligned} \tag{8}$$

where $\tilde{\mathbf{n}}(k)$ is a Gaussian noise with zero mean and covariance $\Lambda = \sigma^2((\mathbf{H}^{(l)})^H \mathbf{H}^{(l)})^{-1}$, $\tilde{\mathbf{y}}(k) = ((\mathbf{H}^{(l)})^H \mathbf{H}^{(l)})^{-1} (\mathbf{H}^{(l)})^H \mathbf{y}(k)$ and the vector $\mathbf{e}_k$ is a column vector whose elements are all zeroes, but the $k$-th which is 1.

Now one models the distribution of $\hat{\mathbf{n}}_j(k)$ as a moment matched Gaussian distribution. One can then calculate an approximated joint symbol probability $\tilde{p}(x_1^{(m,l)}(k),\ldots,x_{j-1}^{(m,l)}(k),x_j(k)|\mathbf{y}(k),\mathbf{H}^{(l)})$ of all the $MN$ possible symbol combinations for $m=1,2,\ldots,M$ and $x_j(k)\in A$:

$$\begin{aligned}
&\tilde{p}\Big(x_1^{(m,l)}(k),\ldots,x_{j-1}^{(m,l)}(k),x_j(k)|\mathbf{y}(k),\mathbf{H}^{(l)}\Big)\\
&\propto \tilde{p}\Big(\mathbf{y}(k)|x_1^{(m,l)}(k),\ldots,x_{j-1}^{(m,l)}(k),x_j(k),\mathbf{H}^{(l)}\Big)\\
&\qquad\qquad p(x_j(k))\prod_{q=1}^{j-1}p(x_q^{(m,l)}(k))\\
&\approx \exp\Big(-\mathbf{w}^H\big(\mathbf{\Pi}_j^{(l)}(k)\big)^{-1}\mathbf{w}\Big)p(x_j(k))\prod_{q=1}^{j-1}p(x_q^{(l,m)}(k))\\
&\stackrel{\text{def}}{=}\psi_m^{(l)}(x_j(k))
\end{aligned}\tag{9}$$

with

$$\mathbf{w}=\tilde{\mathbf{y}}(k)-[x_1^{(l,m)}(k),\ldots,x_{j-1}^{(l,m)}(k),x_j(k),0,\ldots,0]^T,$$

$$\mathbf{\Pi}_j^{(l)}(k)=\mathbf{\Lambda}+\gamma\sum_{q=j+1}^{N_T}\mathbf{e}_q\mathbf{e}_q^T,$$

$$\gamma=\frac{1}{N}\sum_s|a_s|^2$$

and $p(x_j(k))$ is the prior information.

Then $M$ symbol combinations with the largest $\psi_m^{(l)}(x_j(k))$ are selected among the $MN$ possible symbol combinations, resulting in a new set $\Theta_j^{(l)}(k)$.

### B. SGA algorithm summary

To sum up, the SGA algorithm works as follows with channel $\mathbf{H}^{(l)}$ estimated for $k$-th time instant:

1) Initialization: compute $\gamma=\frac{1}{N}\sum_s|a_s|^2$ and $\mathbf{\Pi}_j^{(l)}(k)$ for $j=1,\ldots,N_T$.
2) For each time instant $k$, compute the zero-forcing estimate $\tilde{\mathbf{y}}(k)$ and set $\Theta_0^{(l)}(k)=\emptyset$,
3) The $M$ most significant symbol combinations selection for time instant $k$. For $j=1:N_T$
   a) Compute $\psi_m^{(l)}(x_j(k))$ for all the elements in $\Theta_{j-1}^{(l)}(k)$ and $x_j(k)\in A$ according to Eq. (9). Note that $\Theta_0^{(l)}(k)=\emptyset$ when $j=1$, we only need to compute $N$ possible $\psi_0^{(l)}(x_1(k))$ for $x_1(k)\in A$,
   b) Select the $\min(M,N^j)$ symbol combinations which have the largest $\psi_m^{(l)}(x_j(k))$ and form the set $\Theta_j^{(l)}(k)$.
4) Compute the marginal symbol probabilities for antenna $j=1,2,\ldots,N_T$:
   a) For $m=1,\ldots,M$ and $x_j(k)\in A$, compute

$$\begin{aligned}\phi_m^{(l)}(x_j(k))&=\exp(-(\mathbf{H}^{(l)}\mathbf{v})^H\mathbf{H}^{(l)}\mathbf{v}/\sigma^2)\\
&\qquad p(x_j(k))\prod_{i\neq j}p(x_i^{(l,m)}(k)),\end{aligned}\tag{10}$$

$$\mathbf{v}=\tilde{\mathbf{y}}(k)-[x_1^{(l,m)}(k),\ldots,x_j(k),\ldots x_{N_T}^{(l,m)}(k)]^T$$

   b) Compute the symbol probabilities for $x_j(k)\in A$,

$$\begin{aligned}&\tilde{p}(x_j(k)|\mathbf{y}(k),\mathbf{H}^{(l)})=\\
&\qquad\sum_m\phi_m^{(l)}(x_j(k))/\sum_{x_j(k)}\sum_m\phi_m^{(l)}(x_j(k)).\end{aligned}\tag{11}$$

### C. Mean and covariance computation

The mean $\bar{x}_j^{(l)}(k)$ for $j=1,\ldots,N_T$ can then be computed via Eq. (5).

At the end of the Step 3, we have selected the $M$ most significant symbol combinations

$$\mathbf{x}^{(l,m)}(k)\stackrel{\text{def}}{=}[x_1^{(l,m)}(k),\ldots,x_{N_T}^{(l,m)}(k)]^T$$

and computed

$$\begin{aligned}\varphi_m^{(l)}(k)&\stackrel{\text{def}}{=}p\Big(\mathbf{y}(k)|\mathbf{x}^{(l,m)}(k),\mathbf{H}^{(l)}\Big)\prod_j p(x_j^{(l,m)}(k))\\
&=\exp\Bigg(-\Big(\mathbf{H}^{(l)}\Big(\tilde{\mathbf{y}}(k)-\mathbf{x}^{(l,m)}(k)\Big)\Big)^H\mathbf{H}^{(l)}\\
&\qquad\Big(\tilde{\mathbf{y}}(k)-\mathbf{x}^{(l,m)}(k)\Big)/\sigma^2\Bigg)\prod_j p(x_j^{(l,m)}(k))\end{aligned}\tag{12}$$

for $m=1,\ldots,M$.

Then we can approximate the $(i,j)$-th element of covariance matrix $\mathcal{S}^{(l)}(k)$ as follows:

$$\begin{aligned}\mathcal{S}_{(i,j)}^{(l)}(k)&\approx\frac{1}{\mathcal{Z}(k)}\sum_{m=1}^{M}\varphi_m^{(l)}(k)\\
&\Big(x_i^{(l,m)}(k)-\bar{x}_i^{(l)}(k)\Big)^H\Big(x_j^{(l,m)}(k)-\bar{x}_j^{(l)}(k)\Big)\end{aligned}\tag{13}$$

where $\mathcal{Z}(k)=\sum_m\varphi_m^{(l)}(k)$ is a normalizing constant.

## V. Algorithm Summary

To sum up, the SGA symbol detection based EM algorithm (SGAEM) will work as follows with initial channel estimation $\mathbf{H}^{(0)}$:

1) For $k=1,\ldots,K$:
   a) Run the SGA algorithm to compute the marginal symbol probabilities $\tilde{p}\left(x_j(k)|\mathbf{y}(k),\bar{\mathbf{H}}^{(l)}\right)$ and select $M$ most significant symbol combinations $\mathbf{x}^{(l,m)}(k)$ and compute $\varphi_m^{(l)}(k)$ for $m=1,\ldots,M$.
   b) Compute $\bar{x}_j^{(l)}(k)$ via Eq. (5) and $\mathcal{S}_{(i,j)}^{(l)}(k)$ via Eq. (13) for $i,j=1,\ldots,N_T$.
2) Compute $\mathbf{H}^{l+1}$ via Eq. (7)
3) Repeat the last two steps until a certain number of iterations is reached.

## VI. SIMULATION RESULTS

In this section, we illustrate the performance of SGA symbol detection ($M = 20$) based EM algorithm (SGAEM) via computer simulations. Fig. shows the structure of the SGAEM algorithm for simulation.

To demonstrate the advantages of the proposed SGAEM algorithm which approximate the covariance based on the $M$ most significant symbol combinations, we also present the simulation results of A Posterior Probability (APP) symbol detector based EM algorithm with diagonal approximation of covariance matrix (APPEM-Diag) which is used for parameter estimation in CDMA system [7] [8]. In the E step of APPEM-Diag algorithm, the marginal symbol probabilities are computed with the APP symbol detector and the joint symbol probabilities are approximated as a product of marginal symbol probabilities :

$$\begin{aligned} & p\left(x_i(k), x_j(k)|\mathbf{H}^{(l)}, \mathbf{y}(k)\right) \\ & \approx p\left(x_i(k)|\mathbf{H}^{(l)}, \mathbf{y}(k)\right) p\left(x_j(k)|\mathbf{H}^{(l)}, \mathbf{y}(k)\right) \end{aligned} \tag{14}$$

The covariance matrix $\mathcal{S}^{(l)}(k)$ is then approximated as a diagonal one:

$$\mathcal{S}^{(l)}(k) \approx \mathbf{Diag}\left(\gamma_1^{(l)}(k), \ldots, \gamma_{N_T}^{(l)}(k)\right) \tag{15}$$

$$\gamma_j^{(l)}(k) = \sum_{x_j(k) \in A} |x_j(k) - \bar{x}_j^{(l)}(k)|^2 p\left(x_j(k)|\mathbf{H}^{(l)}, \mathbf{y}(k)\right).$$

In the M step of APPEM-Diag algorithm, the channels are estimated via Eq. (7).

In all our simulations, we set $N_T = N_R = 4$ and consider a 16QAM modulation ($N = 16$). The SNR is defined as $E\{||\mathbf{Hx}||^2\}/E\{||\mathbf{n}||^2\} = \gamma N_T/\sigma^2$. For each SNR, we randomly generate $10^4$ channels realizations. The initial channel estimation $\mathbf{H}^{(0)}$ is computed from the training sequence as follows:

$$\mathbf{H}^{(0)} = \left(\mathbf{X}(0)^H \mathbf{X}(0)\right)^{-1} \mathbf{X}(0)^H \mathbf{Y}(0) \tag{16}$$

where $\mathbf{X}(0)$ is a $4 \times 4$ orthogonal training sequence known to the receiver and $\mathbf{Y}(0)$ is the observation matrix at receiver.

### A. Performance for an uncoded system

In this simulation, each burst length is set to 576 symbols. The symbol error rate (SER) and the channel estimation mean square error (MSE) are shown in Fig. 2 and Fig. 3 respectively. It is seen that the channel estimation quality is improved with EM iterations and thus better SER can be achieved with both algorithms. The SER performance can approach the known channel performance bound APP detector with known channel (APPKnowChan)) with 4 iterations. In the first iteration, the SER performance of SGAEM algorithm is the same as that of APPEM-Diag algorithm.

It can be seen from Fig. 3 that the channel estimation MSE of SGAEM is much lower than that of APPEM-Diag because SGAEM can provide a better approximation to the covariance matrix. After the first iteration, channel MSE performance of SGAEM algorithm is at least 1dB better than that of APPEM-Diag algorithm in low to media SNR levels.

### B. Soft output quality comparison

In this simulation, we will compare the soft output quality of SGAEM algorithm with APPEM-Diag alogirhtm.

A rate $1/2$ Turbo coder with generators 7 and 5 in octal notation is used at the transmitter and a 4-iteration BCJR channel decoder is used at the receiver. The burst length is 1152 bits before channel coding and 238 symbols.

Fig.4 shows the coded BER performance of the APP-KnowChan, SGAEM and APPEM-Diag. It is seen that the BER performance of SGAEM is nearly the same as that of APPEM-Diag in the first iteration when they both use the initial channel estimation $H^{(0)}$ and their performance are both improved with EM iterations. However, the coded BER performance of SGAEM is much better than that of APPEM-Diag because SGAEM can provide a better channel estimation results. After the first iteration, BER performance of SGAEM algorithm is at least 1dB better than that of APPEM-Diag algorithm in low to media SNR levels.

It is also noticed that significant performance improvement can be seen in the first 2-3 iterations of SGAEM algorithm, but the improvement is very small after 3 iterations.

### C. Complexity comparison

The complexity of APP detector based EM algorithm (APPEM-Diag) mentioned in this section grows exponentially the number of transmit antennas ($\mathcal{O}(N^{N_T})$) which is infeasible for practical applications for large systems. However, the complexity of SGAEM proposed in this paper is mainly dominated by the SGA detector where its complexity is only $\mathcal{O}({N_T}^3)$ as discussed in [12].

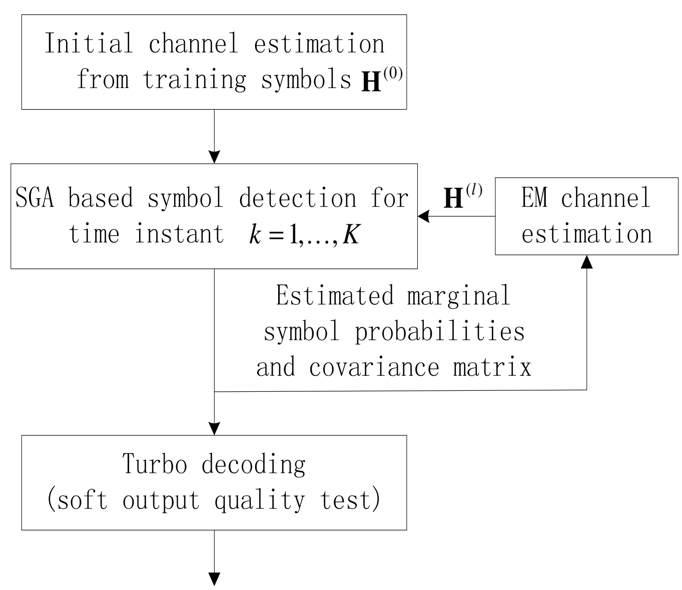


Fig. 1. Illustration of SGAEM algorithm.

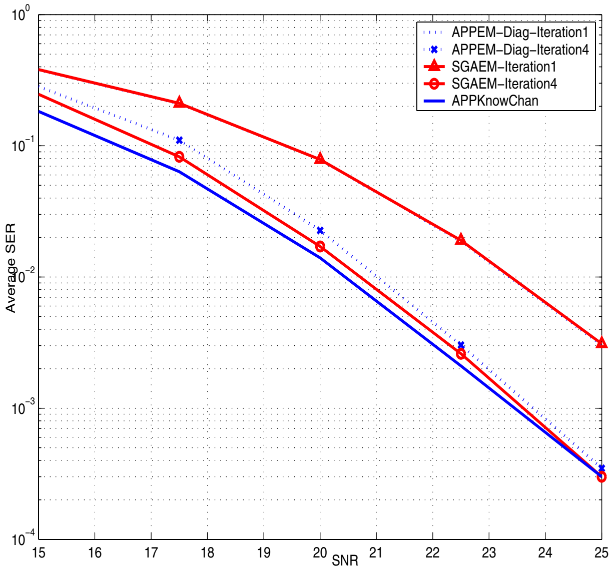


Fig. 2. Uncoded SER performance, $N_T = N_R = 4$, 16QAM.

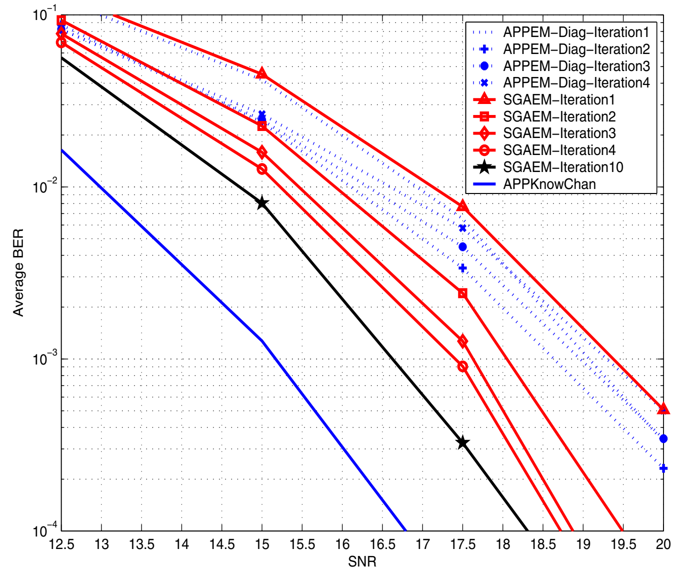


Fig. 4. Coded BER performance comparison.

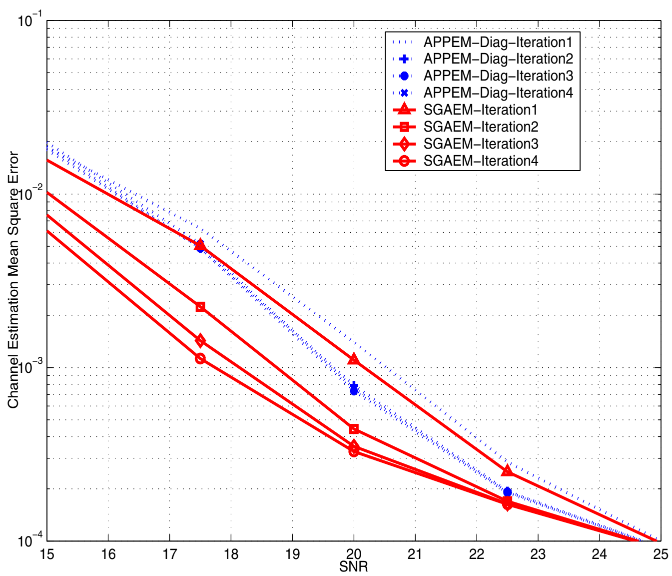


Fig. 3. Channel estimation mean square error, $N_T = N_R = 4$, 16QAM.

## VII. Conclusions

This paper investigates the joint channel estimation and symbol detection for MIMO systems via EM algorithm. In the E-step of the EM algorithm, instead of approximating the posterior covariance of transmitted symbols as a diagonal matrix, all the elements of the covariance are computed based on the $M$ most significant symbol combinations identified via the SGA algorithm. Simulation results show that the proposed algorithm can approach the performance bound.

## Acknowledgements

The authors would like to acknowledge the fruitful discussions with the researchers at Toshiba Research Europe Ltd (Telecommunications Research Laboratory, Bristol) and the support of its directors. The first author would like to thank Toshiba Research Europe Ltd (Telecommunications Research Laboratory, Bristol) for supporting his PHD study at University of Bristol.

## References

[1] G. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J*, pp. 41–59, Autumn 1996.

[2] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum-likelihood from incomplete data via the em algorithm," *J. Roy. Statist. Soc., ser. B*, vol. 39, no. 1, pp. 1–38, 1977.

[3] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Processing Mag.*, pp. 47–59, Nov 1996.

[4] H. V. Poor, "On parameter estimation in DS/SSMA formats," in *Advance in Communications and Signal Processing*. New York: Springer Verlag, 1988, pp. 59–70.

[5] C. N. Georghiades and J. C. Han, "Sequence estimation in the presence of random parameter via EM algorithm," *IEEE Trans. Commun.*, vol. 45, pp. 300–308, Mar 1997.

[6] A. Logothetis and C. Carlemalm, "SAGE algorithm for multipath detection and parameter estimation in asynchronous CDMA systems," *IEEE Trans. Signal Processing*, vol. 48, pp. 3162–3174, Aug. 2000.

[7] M. Kobayashi, J. Boutros, and G. Carie, "Successive inference cancellation with SISO decoding and EM channel estimation," *IEEE J. Select. Areas Commun.*, vol. 19, no. 8, pp. 1450–1460, Aug 2001.

[8] A. Kocian and B. Fleury, "EM-based joint data detection and channel estimation of DS-CDMA signals," *IEEE Trans. Commun.*, vol. 51, no. 10, pp. 1709–1720, Oct 2003.

[9] C. Cozzo and B. L. Hughes, "Joint channel estimation and data detection in space-time communications," *IEEE Trans. Wireless Commun.*, vol. 51, no. 8, pp. 1266–1270, Aug 2003.

[10] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short lengths in a lattice, including a complexity analysis," *Mathematics of Computation*, vol. 44, no. 3, pp. 463–471, Apr. 1985.

[11] J. Luo, K. Pattipati, P. Willett, and F. Hasegawa, "Near-optimal multiuser detection in synchronous CDMA using probabilistic data association," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 361–363, Sept. 2001.

[12] Y. Jia, C. Andrieu, R. J. Piechocki, and M. Sandell, "Gaussian approximation based mixture reduction for near optimal detection for MIMO system," *IEEE Commun. Lett.*, vol. 9, no. 11, pp. 997–999, Nov 2005.